

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Coleman, Roger
Stuart, Susan G.
- (ii) TITLE OF THE INVENTION: A NOVEL HUMAN JAK2 KINASE
- (iii) NUMBER OF SEQUENCES: 5
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 - (B) STREET: 3174 Porter Drive
 - (C) CITY: Palo Alto
 - (D) STATE: CA
 - (E) COUNTRY: US
 - (F) ZIP: 94304
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: FastSEQ Version 1.5
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/567,508
 - (B) FILING DATE: 05-DEC-1995
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Billings, Lucy J.
 - (B) REGISTRATION NUMBER: 36,749
 - (C) REFERENCE/DOCKET NUMBER: PF-0049US
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 650-855-0555
 - (B) TELEFAX: 650-845-4166

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4482 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: Placenta
 - (B) CLONE: 179527

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCACGCGTCC	GGTTGCCAAC	CCGCAGGCGA	CTGGGCGCTT	CATCCCACCC	TCACCCCTTT	60
CCAGCCAAGG	TGGCTGATCG	GAGTCAGGCT	CTCGAGGTCG	CATTGCCACG	AAACGGNGTG	120
TGTGAGCGCG	TTGTCCCCGG	NCCCCGGGGC	CACTTCCCCT	CGGCCTAGNA	GA CTGGACTG	180
GGGAAGGACG	GGTCTGTTGT	ACCCGGGAGG	TGGAAGGAAA	AGCCGAAAGC	GGAGAAGTGT	240
GCGGGAGGGG	AGTCTCCGCG	CGGAGGNAGA	CCGGNCTCCT	CCAGTGCAGG	TTGTGCGCTG	300
GGGAGCCAGC	CASGGCAAAT	GTTCTGAAAA	AGACTCTGCA	TGGGAATGGC	CTGCCTTACG	360
ATGACAGAAA	TGGAGGGAAC	ATCCACCTCT	TCTATATATC	AGAATGGTGA	TATTTCTGGA	420
AATGCCAATT	CTATGAAGCA	AATAGATCCA	GTTCTTCAGG	TGTATCTTTA	CCATTCCCTT	480
GGGAAATCTG	AGGCAGATTA	TCTGACCTTT	CCATCTGGGG	AGTATGTTGG	AGAAGAAATC	540
TGTATTGCTG	CTTCTAAAGC	TTGTGGTATC	ACACCTGTGT	ATCATAATAT	GTTTGCTTTA	600
ATGAGTGAAA	CAGAAAGGAT	CTGGTATCCA	CCCAACCATG	TCTTCCATAT	AGATGAGTCA	660
ACCGGCATA	ATGTACTCTA	CAGAATAAGA	TTTTACTTTT	CTCGTTGGTA	TTGCAGTGGC	720
AGCAACAGAG	CCTATCGGCA	TGGAATATCT	CGAGGTGCTG	AAGCTCCTCT	TCTTGATGAC	780
TTTGTTCATG	CTTACCTCTT	TGCTCAGTGG	CGGCATGATT	TTGTGCATGG	ATGGATAAAA	840
GTACCTGTGA	CTCATGAAAC	ACAGGAAGAA	TGTCTTGGGA	TGACAGTGTT	AGATATGATG	900
AGAATAGCCA	AAGAAAACGA	TCAAACCCCA	CTGGCCATCT	ATAACTCTAT	CAGCTACAAG	960
CAATTCTTAC	CACAATGTAT	TCGAGCAAAG	ATCCAAGACT	ATCATATTTT	GACAAGGAAG	1020
CGAATAAGGT	ACAGATTTTG	CAGATTTATT	CAGCAATTCA	GCCAATGCAA	AGCCACTGCC	1080
AGAAACTTGA	AACTTAAAGT	TCTTATAAAT	CTGGAAACTC	TGCAGTCTGC	CTTCTACACA	1140
GAGAAATTTG	AAGTAAAAGA	ACCTGGAAGT	GGTCCTTCAG	GTGAGGAGAT	TTTGCAACC	1200
ATTATAATAA	CTGGAAACGG	TGGAATTCAG	TGGTCAAGAG	GGAAACATAA	AGAAAGTGAG	1260
ACACTGACAG	AACAGGATTT	ACAGTTATAT	TGCGATTTTC	CTAATATTAT	TGATGTCAGT	1320
ATTAAGCAAG	CAAACCAAGA	GGGTTCAAAT	GAAAGCCGAG	TTGTAACAT	CCATAAGCAA	1380
GATGGTAAAA	ATCTGGAAAT	TGAACTTAGC	TCATTAAGGG	AAGCTTTGTC	TTTCGTGTCA	1440
TTAATTGATG	GATATTATAG	ATTAAC TGCA	GATGCACATC	ATTACCTCTG	TAAAGAAGTA	1500
GCACCTCCAG	CCGTGCTTGA	AAATATACAA	AGCAACTGTC	ATGGCCCAAT	TTCGATGGAT	1560
TTTGCCATTA	GTAAACTGAA	GAAAGCAGGT	AATCAGACTG	GACTGTATGT	ACTTCGATGC	1620
AGTCCTAAGG	ACTTTAATAA	ATATTTTTTG	ACTTTTGCTG	TCGAGCGAGA	AAATGTCATT	1680
GAATATAAAC	ACTGTTTGAT	TACAAAAAAT	GAGAATGAAG	AGTACAACCT	CAGTGGGACA	1740
AAGAAGAACT	TCAGCAGTCT	TAAAGATCCT	TTGAATTGTT	ACCAGATGGA	AACTGTTTCG	1800
TCAGACAATA	TAATTTTCCA	GTTTACTAAA	TGCTGTCCCC	CAAAGCCAAA	AGATAAATCA	1860
AACCTTCTAG	TCTTCAGAAC	GAATGGTGTT	TCTGATGTAC	CAACCTCACC	AACATTACAG	1920
AGGCCTACTC	ATATGAACCA	AATGGTGTTT	CACAAAATCA	GAAATGAAGA	TTTGATATTT	1980
AATGAAAGCC	TTGGCCAAGG	CACTTTTTACA	AAGATTTTTA	AAGGCGTACG	AAGAGAAGTA	2040
GGAGACTACG	GTCAACTGCA	TGAAACAGAA	GTTCTTTTAA	AAGTTCTGGA	TAAAGCACAC	2100
AGGAACTATT	CAGAGTCTTT	CTTTGAAGCA	GCAAGTATGA	TGAGCAAGCT	TTCTCAGCAG	2160
CATTTGGTTT	TAAATTATGG	AGTATGTGTC	TGTGGAGACG	AGAATATTCT	GGTTCAGGAG	2220
TTTGTAATAA	TTGGATCACT	AGATACATAT	CTGAAAAAGA	ATAAAAATTG	TATAAATATA	2280
TTATGGAAAC	TTGAAGTTGC	TAAACAGTTG	GCATGGGCCA	TGCATTTTCT	AGAAGAAAAC	2340
ACCCTTATTC	ATGGGAATGT	ATGTGCCAAA	AATATTCTGC	TTATCAGAGA	AGAAGACAGG	2400
AAGACAGGAA	ATCCTCCTTT	ATCCAAACTT	AGTGATCCTG	GCATTAGTAT	TACAGTTTTG	2460
CCAAAGGACA	TTCTTCAGGA	GAGAATACCA	TGGGTACCAC	CTGAATGCAT	TGAAAATCCT	2520
AAAAATTTAA	ATTTGGCAAC	AGACAAATGG	AGTTTTGGTA	CCACTTTGTG	GGAAATCTGC	2580
AGTGGAGGAG	ATAAACCTCT	AAGTGCTCTG	GATTCTCAAA	GAAAGCTACA	ATTTTATGAA	2640
GATAGGCATC	AGCTTCCTGC	ACCAAAGTGG	GCAGAATTAG	CAAACCTTAT	AAATAATTGT	2700
ATGGATTATG	AACCAGATTT	CAGGCCTTCT	TTCAGAGCCA	TCATACGAGA	TCTTAACAGT	2760
TTGTTTACTC	CAGATTATGA	ACTATTAAAC	GAAAATGACA	TGTTACCAA	TATGAGGATA	2820
GGTGCCTTGG	GGTTTTCTGG	TGCCTTTGAA	GACCGGGATC	CTACACAGTT	TGAAGAGAGA	2880
CATTTGAAAT	TTCTACAGCA	ACTTGGCAAG	GGTAATTTTG	GGAGTGTGGA	GATGTGCCGG	2940
TATGACCCTC	TACAGGACAA	CACTGGGGAG	GTGGTCGCTG	TAAAAAAGCT	TCAGCATAGT	3000
ACTGAAGAGC	ACCTAAGAGA	CTTTGAAAGG	GAAATTGAAA	TCCTGAAATC	CCTACAGCAT	3060
GACAACATTT	TAAAGTACAA	GGGAGTGTGC	TACAGTGCTG	GTCGGCGTAA	TCTAAAATTA	3120
ATTATGGAAT	ATTTACCATA	TGGAAGTTTA	CGAGACTATC	TTCAAAAACA	TAAAGAACGG	3180
ATAGATCACA	TAAAACCTCT	GCAGTACACA	TCTCAGATAT	GCAAGGGTAT	GGAGTATCTT	3240
GGTACAAAAA	GGTATATCCA	CAGGGATCTG	GCAACGAGAA	ATATATTGGT	GGAGAACGAG	3300
AACAGAGTTA	AAATTGGRGA	TTTTGGGTTA	ACCAAAGTCT	TGCCACAAGA	CAAAGAATAC	3360
TATAAAGTAA	AAGAACCTGG	TGAAAGTCCC	ATATTCTGGT	ATGCTCCAGA	ATCACTGACA	3420
GAGAGCAAGT	TTTCTGTGGC	CTCAGATGTT	TGGAGCTTTG	GAGTGGTTCT	GTATGAACTT	3480
TTACACATA	TTGAGAAAG	TAAAAGTCCA	CCAGCGGAAT	TTATGCGTAT	GATTGGCAAT	3540
GACAAACAAG	GACAGATGAT	CGTGTTCAT	TTGATAGAAC	TTTTGAAGAA	TAATGGAAGA	3600
TTACCAAGAC	CAGATGGATG	CCCAGATGAG	ATCTATATGA	TCATGACAGA	ATGCTGGAAC	3660
AATAATGTAA	ATCAACGCCC	CTCCTTTAGG	GATCTAGCTC	TTCGAGTGGA	TCAAATAAGG	3720
GATAACATGG	CTGGATGAAA	GAAATGACCT	TCATTCTGAG	ACCAAAGTAG	ATTTACAGAA	3780
CAAAGTTTTA	TATTTACAT	TGCTGTGGAC	TATTATTACA	TATATCATTA	TTATATAAAT	3840
CATGATGCTA	GCCAGCAAAAG	ATGTGAAAAT	ATCTGCTCAA	AACTTTCAAA	GTTTAGTAAG	3900

TTTTTCTTCA	TGAGGCCACC	AGTAAAAGAC	ATTAATGAGA	ATTCCTTAGC	AAGGATTTTG	3960
TAAGAAGTTT	CTTAAACATT	GTCAGTTAAC	ATCACTCTTG	TCTGGCAAAA	GAAAAAAAAAT	4020
AGACTTTTTC	AACTCAGCTT	TTTGAGACCT	GAAARAATTA	TTATGTAAAT	TTTGCAATGT	4080
TAAAGATGCA	CAGAATATGT	ATGTATAGTT	TTTACCACAG	TGGATGTATA	ATACCTTGGC	4140
ATCTTGTGTG	ATGTTTAACA	CACATGAGGG	CTGGTGTTCA	TTAATACTGT	TTTCTAATTT	4200
TTCCATGGTT	AATCTATAAT	TAATTACTTC	ACTAAACAAA	CAAATTAAGA	TGTTTCAGATA	4260
ATTGAATAAG	TACCTTTGTG	TCCTTGTTCA	TTTATATCGC	TGGCCAGCAT	TATAAGCAGG	4320
TGTATACTTT	TAGCTTGTAG	TTCCATGTAC	TGTAAATATT	TTTCACATAA	AGGGAACAAA	4380
TGTCTAGTTT	TATTTGTATA	GGAAATTTGC	CCTGACCCTA	AATAATACAT	TTTGAAATGA	4440
AACAAGCTTA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	AG		4482

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1132 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY:
- (B) CLONE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Gly	Met	Ala	Cys	Leu	Thr	Met	Thr	Glu	Met	Glu	Gly	Thr	Ser	Thr	1	5	10	15
Ser	Ser	Ile	Tyr	Gln	Asn	Gly	Asp	Ile	Ser	Gly	Asn	Ala	Asn	Ser	Met	20	25	30	
Lys	Gln	Ile	Asp	Pro	Val	Leu	Gln	Val	Tyr	Leu	Tyr	His	Ser	Leu	Gly	35	40	45	
Lys	Ser	Glu	Ala	Asp	Tyr	Leu	Thr	Phe	Pro	Ser	Gly	Glu	Tyr	Val	Gly	50	55	60	
Glu	Glu	Ile	Cys	Ile	Ala	Ala	Ser	Lys	Ala	Cys	Gly	Ile	Thr	Pro	Val	65	70	75	80
Tyr	His	Asn	Met	Phe	Ala	Leu	Met	Ser	Glu	Thr	Glu	Arg	Ile	Trp	Tyr	85	90	95	
Pro	Pro	Asn	His	Val	Phe	His	Ile	Asp	Glu	Ser	Thr	Arg	His	Asn	Val	100	105	110	
Leu	Tyr	Arg	Ile	Arg	Phe	Tyr	Phe	Pro	Arg	Trp	Tyr	Cys	Ser	Gly	Ser	115	120	125	
Asn	Arg	Ala	Tyr	Arg	His	Gly	Ile	Ser	Arg	Gly	Ala	Glu	Ala	Pro	Leu	130	135	140	
Leu	Asp	Asp	Phe	Val	Met	Ser	Tyr	Leu	Phe	Ala	Gln	Trp	Arg	His	Asp	145	150	155	160
Phe	Val	His	Gly	Trp	Ile	Lys	Val	Pro	Val	Thr	His	Glu	Thr	Gln	Glu	165	170	175	
Glu	Cys	Leu	Gly	Met	Thr	Val	Leu	Asp	Met	Met	Arg	Ile	Ala	Lys	Glu	180	185	190	
Asn	Asp	Gln	Thr	Pro	Leu	Ala	Ile	Tyr	Asn	Ser	Ile	Ser	Tyr	Lys	Thr	195	200	205	
Phe	Leu	Pro	Gln	Cys	Ile	Arg	Ala	Lys	Ile	Gln	Asp	Tyr	His	Ile	Leu	210	215	220	
Thr	Arg	Lys	Arg	Ile	Arg	Tyr	Arg	Phe	Arg	Arg	Phe	Ile	Gln	Gln	Phe	225	230	235	240
Ser	Gln	Cys	Lys	Ala	Thr	Ala	Arg	Asn	Leu	Lys	Leu	Lys	Tyr	Leu	Ile	245	250	255	
Asn	Leu	Glu	Thr	Leu	Gln	Ser	Ala	Phe	Tyr	Thr	Glu	Lys	Phe	Glu	Val	260	265	270	
Lys	Glu	Pro	Gly	Ser	Gly	Pro	Ser	Gly	Glu	Glu	Ile	Phe	Ala	Thr	Ile	275	280	285	
Ile	Ile	Thr	Gly	Asn	Gly	Gly	Ile	Gln	Trp	Ser	Arg	Gly	Lys	His	Lys	290	295	300	
Glu	Ser	Glu	Thr	Leu	Thr	Glu	Gln	Asp	Leu	Gln	Leu	Tyr	Cys	Asp	Phe	305	310	315	320
Pro	Asn	Ile	Ile	Asp	Val	Ser	Ile	Lys	Gln	Ala	Asn	Gln	Glu	Gly	Ser	325	330	335	
Asn	Glu	Ser	Arg	Val	Val	Thr	Ile	His	Lys	Gln	Asp	Gly	Lys	Asn	Leu	340	345	350	
Glu	Ile	Glu	Leu	Ser	Ser	Leu	Arg	Glu	Ala	Leu	Ser	Phe	Val	Ser	Leu	355	360	365	
Ile	Asp	Gly	Tyr	Tyr	Arg	Leu	Thr	Ala	Asp	Ala	His	His	Tyr	Leu	Cys	370	375	380	

Lys	Glu	Val	Ala	Pro	Pro	Ala	Val	Leu	Glu	Asn	Ile	Gln	Ser	Asn	Cys
385					390					395					400
His	Gly	Pro	Ile	Ser	Met	Asp	Phe	Ala	Ile	Ser	Lys	Leu	Lys	Lys	Ala
				405					410					415	
Gly	Asn	Gln	Thr	Gly	Leu	Tyr	Val	Leu	Arg	Cys	Ser	Pro	Lys	Asp	Phe
			420					425					430		
Asn	Lys	Tyr	Phe	Leu	Thr	Phe	Ala	Val	Glu	Arg	Glu	Asn	Val	Ile	Glu
		435					440					445			
Tyr	Lys	His	Cys	Leu	Ile	Thr	Lys	Asn	Glu	Asn	Glu	Glu	Tyr	Asn	Leu
	450					455					460				
Ser	Gly	Thr	Lys	Lys	Asn	Phe	Ser	Ser	Leu	Lys	Asp	Leu	Leu	Asn	Cys
465					470					475					480
Tyr	Gln	Met	Glu	Thr	Val	Arg	Ser	Asp	Asn	Ile	Ile	Phe	Gln	Phe	Thr
				485					490					495	
Lys	Cys	Cys	Pro	Pro	Lys	Pro	Lys	Asp	Lys	Ser	Asn	Leu	Leu	Val	Phe
			500					505					510		
Arg	Thr	Asn	Gly	Val	Ser	Asp	Val	Pro	Thr	Ser	Pro	Thr	Leu	Gln	Arg
		515					520					525			
Pro	Thr	His	Met	Asn	Gln	Met	Val	Phe	His	Lys	Ile	Arg	Asn	Glu	Asp
	530					535					540				
Leu	Ile	Phe	Asn	Glu	Ser	Leu	Gly	Gln	Gly	Thr	Phe	Thr	Lys	Ile	Phe
545					550					555					560
Lys	Gly	Val	Arg	Arg	Glu	Val	Gly	Asp	Tyr	Gly	Gln	Leu	His	Glu	Thr
				565					570					575	
Glu	Val	Leu	Leu	Lys	Val	Leu	Asp	Lys	Ala	His	Arg	Asn	Tyr	Ser	Glu
			580					585					590		
Ser	Phe	Phe	Glu	Ala	Ala	Ser	Met	Met	Ser	Lys	Leu	Ser	His	Lys	His
		595					600					605			
Leu	Val	Leu	Asn	Tyr	Gly	Val	Cys	Val	Cys	Gly	Asp	Glu	Asn	Ile	Leu
	610					615					620				
Val	Gln	Glu	Phe	Val	Lys	Phe	Gly	Ser	Leu	Asp	Thr	Tyr	Leu	Lys	Lys
625					630					635					640
Asn	Lys	Asn	Cys	Ile	Asn	Ile	Leu	Trp	Lys	Leu	Glu	Val	Ala	Lys	Gln
				645					650					655	
Leu	Ala	Trp	Ala	Met	His	Phe	Leu	Glu	Glu	Asn	Thr	Leu	Ile	His	Gly
			660					665					670		
Asn	Val	Cys	Ala	Lys	Asn	Ile	Leu	Leu	Ile	Arg	Glu	Glu	Asp	Arg	Lys
		675					680					685			
Thr	Gly	Asn	Pro	Pro	Phe	Ile	Lys	Leu	Ser	Asp	Pro	Gly	Ile	Ser	Ile
	690					695					700				
Thr	Val	Leu	Pro	Lys	Asp	Ile	Leu	Gln	Glu	Arg	Ile	Pro	Trp	Val	Pro
705					710					715					720
Pro	Glu	Cys	Ile	Glu	Asn	Pro	Lys	Asn	Leu	Asn	Leu	Ala	Thr	Asp	Lys
				725					730					735	
Trp	Ser	Phe	Gly	Thr	Thr	Leu	Trp	Glu	Ile	Cys	Ser	Gly	Gly	Asp	Lys
			740					745					750		
Pro	Leu	Ser	Ala	Leu	Asp	Ser	Gln	Arg	Lys	Leu	Gln	Phe	Tyr	Glu	Asp
		755					760					765			
Arg	His	Gln	Leu	Pro	Ala	Pro	Lys	Trp	Ala	Glu	Leu	Ala	Asn	Leu	Ile
	770					775					780				
Asn	Asn	Cys	Met	Asp	Tyr	Glu	Pro	Asp	Phe	Arg	Pro	Ser	Phe	Arg	Ala
785					790					795					800
Ile	Ile	Arg	Asp	Leu	Asn	Ser	Leu	Phe	Thr	Pro	Asp	Tyr	Glu	Leu	Leu
				805					810					815	
Thr	Glu	Asn	Asp	Met	Leu	Pro	Asn	Met	Arg	Ile	Gly	Ala	Leu	Gly	Phe
			820					825					830		
Ser	Gly	Ala	Phe	Glu	Asp	Arg	Asp	Pro	Thr	Gln	Phe	Glu	Glu	Arg	His
		835					840					845			
Leu	Lys	Phe	Leu	Gln	Gln	Leu	Gly	Lys	Gly	Asn	Phe	Gly	Ser	Val	Glu
	850					855					860				
Met	Cys	Arg	Tyr	Asp	Pro	Leu	Gln	Asp	Asn	Thr	Gly	Glu	Val	Val	Ala
865					870					875					880
Val	Lys	Lys	Leu	Gln	His	Ser	Thr	Glu	Glu	His	Leu	Arg	Asp	Phe	Glu
				885					890					895	
Arg	Glu	Ile	Glu	Ile	Leu	Lys	Ser	Leu	Gln	His	Asp	Asn	Ile	Val	Lys
			900					905					910		
Tyr	Lys	Gly	Val	Cys	Tyr	Ser	Ala	Gly	Arg	Arg	Asn	Leu	Lys	Leu	Ile

[illegible]

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1129 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

Met Gly Met Ala Cys Leu Thr Met Thr Glu Met Glu Ala Thr Ser Thr
 1           5           10           15
Ser Pro Val His Gln Asn Gly Asp Ile Pro Gly Ser Ala Asn Ser Val
          20           25           30
Lys Gln Ile Glu Pro Val Leu Gln Val Tyr Leu Tyr His Ser Leu Gly
          35           40           45
Gln Ala Glu Gly Glu Tyr Leu Lys Phe Pro Ser Gly Glu Tyr Val Ala
          50           55           60
Glu Glu Ile Cys Val Ala Ala Ser Lys Ala Cys Gly Ile Thr Pro Val
 65           70           75           80
Tyr His Asn Met Phe Ala Leu Met Ser Glu Thr Glu Arg Ile Trp Tyr
          85           90           95
Pro Pro Asn His Val Phe His Ile Asp Glu Ser Thr Arg His Asp Ile
          100          105          110
Leu Tyr Arg Ile Arg Phe Tyr Phe Pro His Trp Tyr Cys Ser Gly Ser
          115          120          125
Ser Arg Thr Tyr Arg Tyr Gly Val Ser Arg Gly Ala Glu Ala Pro Leu
          130          135          140
Leu Asp Asp Phe Val Met Ser Tyr Leu Phe Val Gln Trp Arg His Asp
145          150          155          160
Phe Val His Gly Trp Ile Lys Val Pro Val Thr His Glu Thr Gln Glu
          165          170          175
Glu Cys Leu Gly Met Ala Val Leu Asp Met Met Arg Ile Ala Lys Glu
          180          185          190
Lys Asp Gln Thr Pro Leu Ala Val Tyr Asn Ser Val Ser Tyr Lys Thr
          195          200          205
Phe Leu Pro Lys Cys Val Arg Ala Lys Ile Gln Asp Tyr His Ile Leu
          210          215          220
Thr Arg Lys Arg Ile Arg Tyr Arg Phe Arg Arg Phe Ile Gln Gln Phe
225          230          235          240
Ser Gln Cys Lys Ala Thr Ala Arg Asn Leu Lys Leu Lys Tyr Leu Ile
          245          250          255
Asn Leu Glu Thr Leu Gln Ser Ala Phe Tyr Thr Glu Gln Phe Glu Val
          260          265          270
Lys Glu Ser Ala Arg Gly Pro Ser Gly Glu Glu Ile Phe Ala Thr Ile
          275          280          285
Ile Ile Thr Gly Asn Gly Gly Ile Gln Trp Ser Arg Gly Lys His Lys
          290          295          300
Glu Ser Glu Thr Leu Thr Glu Gln Asp Val Gln Leu Tyr Cys Asp Phe
305          310          315          320
Pro Asp Ile Ile Asp Val Ser Ile Lys Gln Ala Asn Gln Glu Cys Ser
          325          330          335
Asn Glu Ser Arg Ile Val Thr Val His Lys Gln Asp Gly Lys Val Leu
          340          345          350
Glu Ile Glu Leu Ser Ser Leu Lys Glu Ala Leu Ser Phe Val Ser Leu
          355          360          365
Ile Asp Gly Tyr Tyr Arg Leu Thr Ala Asp Ala His His Tyr Leu Cys
          370          375          380
Lys Glu Val Ala Pro Pro Ala Val Leu Glu Asn Ile His Ser Asn Cys
385          390          395          400
His Gly Pro Ile Ser Met Asp Phe Ala Ile Ser Lys Leu Lys Lys Ala
          405          410          415
Gly Asn Gln Thr Gly Leu Tyr Val Leu Arg Cys Ser Pro Lys Asp Phe
          420          425          430
Asn Lys Tyr Phe Leu Thr Phe Ala Val Glu Arg Glu Asn Val Ile Glu

```

[illegible]

Leu Ala Thr Arg Asn Ile Leu Val Glu Asn Glu Asn Arg Val Lys Ile
 980 985
 Gly Asp Phe Gly Leu Thr Lys Val Leu Pro Gln Asp Lys Glu Tyr Tyr
 995 1000 1005
 Lys Val Lys Glu Pro Gly Glu Ser Pro Ile Phe Trp Tyr Ala Pro Gln
 1010 1015 1020
 Ser Leu Thr Glu Ser Lys Phe Ser Val Ala Ser Asp Val Trp Ser Phe
 025 1030 1035 1040
 Gly Val Val Leu Tyr Glu Leu Phe Thr Tyr Ile Glu Lys Ser Lys Ser
 1045 1050 1055
 Pro Pro Val Glu Phe Met Arg Met Ile Gly Asn Asp Lys Gln Gly Gln
 1060 1065 1070
 Met Ile Val Phe His Leu Ile Glu Leu Leu Lys Ser Asn Gly Arg Leu
 1075 1080 1085
 Pro Arg Pro Glu Gly Cys Pro Asp Glu Ile Tyr Val Ile Met Thr Glu
 1090 1095 1100
 Cys Trp Asn Asn Asn Val Ser Gln Arg Pro Ser Phe Arg Asp Leu Ser
 105 1110 1115 1120
 Phe Gly Trp Ile Lys Cys Gly Thr Val
 1125

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGGCGGAAGT GCTCTCGGCG GAAG

24

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGTGTGCTAC AGTGCTGGTC GTCG

24